

2009

A new approach to sparse image representation using MMV and K-SVD

Jie Yang

University of Wollongong, jy962@uow.edu.au

Abdesselam Bouzerdoun

University of Wollongong, bouzer@uow.edu.au

Son Lam Phung

University of Wollongong, phung@uow.edu.au

Publication Details

Yang, J., Bouzerdoun, A. & Phung, S. (2009). A new approach to sparse image representation using MMV and K-SVD. In J. Blanc-Talon, W. Philips, D. Popescu & P. Scheunders (Eds.), *Advanced concepts for intelligent vision systems : ACIVS 2009* (pp. 200-209). Berlin, Germany: Springer-Verlag.

A new approach to sparse image representation using MMV and K-SVD

Abstract

This paper addresses the problem of image representation based on a sparse decomposition over a learned dictionary. We propose an improved matching pursuit algorithm for Multiple Measurement Vectors (MMV) and an adaptive algorithm for dictionary learning based on multi-Singular Value Decomposition (SVD), and combine them for image representation. Compared with the traditional K-SVD and orthogonal matching pursuit MMV (OMPMMV) methods, the proposed method runs faster and achieves a higher overall reconstruction accuracy.

Keywords

mmv, approach, representation, image, k, sparse, svd

Disciplines

Physical Sciences and Mathematics

Publication Details

Yang, J., Bouzerdoun, A. & Phung, S. (2009). A new approach to sparse image representation using MMV and K-SVD. In J. Blanc-Talon, W. Philips, D. Popescu & P. Scheunders (Eds.), *Advanced concepts for intelligent vision systems : ACIVS 2009* (pp. 200-209). Berlin, Germany: Springer-Verlag.

A New Approach to Sparse Image Representation Using MMV and K-SVD

Jie Yang, Abdesselam Bouzerdoum, and Son Lam Phung

School of Electrical, Computer and Telecommunications Engineering
University of Wollongong, Wollongong, NSW 2522, Australia

Abstract. This paper addresses the problem of image representation based on a sparse decomposition over a learned dictionary. We propose an improved matching pursuit algorithm for Multiple Measurement Vectors (MMV) and an adaptive algorithm for dictionary learning based on multi-Singular Value Decomposition (SVD), and combine them for image representation. Compared with the traditional K-SVD and *orthogonal matching pursuit* MMV (OMPMMV) methods, the proposed method runs faster and achieves a higher overall reconstruction accuracy.

1 Introduction

Signal representation is important for efficient signal processing, data compression, pattern recognition and machine learning. The problem of how to select a set of basis vectors for efficient representation of signals in a given dataset has been extensively investigated in the past [1, 2]. This problem can be described mathematically as follows: given an original signal \mathbf{y} in an n -dimensional space and a set of basis vectors, find, within a preset tolerance, a compact representation of \mathbf{y} using the subspace spanned by the basis vectors. The development of pursuit algorithms such as *orthogonal matching pursuit* [5] from *compressed sensing* [3, 4], with the capability to find a sparse representation, has offered new approaches for tackling the aforementioned problem. Using an over-complete dictionary D consisting of k basis vectors or *atoms*, the original signal can be decomposed by solving the following system of linear equations:

$$\mathbf{y} = D\mathbf{x} \quad (1)$$

In the past, the K-SVD approach has been proposed for dictionary-based learning for a sparse representation [6, 7, 8]. However, this approach has a number of problems in dealing with high-dimensional signals. First, K-SVD depends heavily on a pursuit algorithm to calculate the sparse coefficients. Second, it updates the dictionary atom-by-atom during each iteration. Furthermore, the K-SVD requires large storage because the computed non-zero coefficients reside in different locations.

To overcome the K-SVD disadvantages, we propose a new approach to signal representation that is based on the concept of Multiple Measurement Vectors (MMV) [9], [10]. MMV aims to find a solution where most nonzero elements

are clustered in a few rows. The proposed method requires significantly less storage compared to K-SVD. Furthermore, it allows the simultaneous update of several atoms, which leads to faster convergence and better reconstruction accuracy. Here, the new method is applied to a simulated data and the problem of image representation and its performance is assessed in terms of reconstruction accuracy and convergence speed.

The paper is organized as follows. Section 2 provides background information on compressed sensing and K-SVD algorithm. Section 3 presents the proposed approach based on Multiple Measurement Vectors. Section 4 gives an analysis of the proposed approach in an image representation task. Section 5 presents concluding remarks.

2 Background

This section reviews the basic concepts of compressed sensing and discusses two existing algorithms, namely *orthogonal matching pursuit*(OMP) and K-SVD.

2.1 Compressed Sensing

Compressed sensing (CS) aims to find a sparse solution to the problem in (1), where most of the elements of the solution vector \mathbf{x} are zero [3, 4]. Compressed sensing algorithms can be divided into two broad categories: (i) Single Measurement Vector (SMV) [5], [11] where the solution is a vector; and (ii) Multiple Measurement Vectors (MMV) [9], [10] where the solution is a two-dimensional array, or matrix.

In SMV, the problem can be formulated as:

$$\text{minimize } \|\mathbf{x}\|_0, \text{ subject to } \mathbf{y} = D\mathbf{x} \quad (2)$$

where $\mathbf{y} \in \mathbb{R}^n$ is an observable (measurement) vector, $D \in \mathbb{R}^{n \times k}$ is a known dictionary containing k basis vectors or atoms, and $\|\mathbf{x}\|_0$ denotes the number of nonzero elements in \mathbf{x} . This problem can be solved by several approaches, including greedy algorithms such as Orthogonal Matching Pursuit [5], and non-convex local optimization such as FOCUSS algorithm [11].

On the other hand, the aim in MMV is to

$$\text{minimize } \|X\|_0, \text{ subject to } Y = DX \quad (3)$$

where $Y \in \mathbb{R}^{n \times m}$ is a matrix comprising multiple measurement vectors, X is the solution matrix ($X \in \mathbb{R}^{k \times m}$), and $\|X\|_0$ denotes the number of non-empty rows in X . A non-empty row has at least one non-zero entry. The OMP technique and its variants developed for SMV have been extended to tackle MMV [9].

In this paper, we focus on the Orthogonal Matching Pursuit for MMV algorithm (OMPMMV) proposed in [10]. The steps of this iterative algorithm are summarized in Table 1. In this algorithm, \mathcal{S}_t denotes the set of selected atoms after the t -th iteration, and R_t represents the residual error, obtained by using \mathcal{S}_t for signal reconstruction.

Table 1. OMPMMV algorithm [10]

<ol style="list-style-type: none"> 1. Initialize $\mathcal{S}_0 = \emptyset$ (empty set) and $R_0 = Y$. 2. Repeat steps 3 to 6 until convergence. 3. Compute $A = D^T R_{t-1}$. 4. Select from D the column (or atom) \mathbf{d}_i^t, which corresponds to the row A_i with the largest magnitude in A. 5. Update the set of selected atoms: $\mathcal{S}_t = \mathcal{S}_{t-1} \cup \mathbf{d}_i^t$. 6. Update the residual $R_t = Y - \mathcal{S}_t X_t$, where $X_t = [\mathcal{S}_t^T \mathcal{S}_t]^{-1} \mathcal{S}_t^T Y$.

2.2 K-SVD Method

Consider an $n \times m$ matrix Y , comprising m measurement vectors, and a known fixed dictionary $D \in \mathbb{R}^{n \times k}$. The problem of finding a compact signal representation can be expressed as

$$\text{minimize } S_m(X), \text{ subject to } Y = DX \tag{4}$$

where $S_m(X)$ is a sparsity measure of X . In [6, 7, 8], $S(X)$ is defined in terms of the l^0 -norm for columns \mathbf{x}_i of X . In K-SVD method, each column \mathbf{y}_i of Y is extracted and the traditional SMV method is applied on the pair (\mathbf{y}_i, D) to obtain a solution in the form of

$$\hat{\mathbf{x}}_i = \underset{\mathbf{x}_i}{\text{argmin}} \|\mathbf{x}_i\|_0, \text{ subject to } \mathbf{y}_i = D\mathbf{x}_i \tag{5}$$

Note that K-SVD assumes that the dictionary D is unknown so it calculates D as well. Hence, the optimization problem in (5) becomes

$$\left(\hat{D}, \hat{X} \right) = \underset{D, X}{\text{argmin}} \{ S_m(X) + \|Y - DX\|_F^2 \}, \tag{6}$$

where $\|\cdot\|_F^2$ denotes the Frobenius norm.

The main advantage of K-SVD is that it not only finds a sparse solution for each column of X , but also updates simultaneously the dictionary D via SVD. The reader is referred to [6, 7, 8] for a wide range of applications of K-SVD.

3 MMV-Based Signal Representation

To describe the proposed MMV-based approach to signal representation, we first give some definitions:

Definition 1. Given a matrix X , its *sparse I rank* $S_1(X)$ is the largest number of non-zero entries in any column of X .

Definition 2. Given a matrix X , its *sparse II rank* $S_2(X)$ is the number of non-empty rows in X . A non-empty row must have at least one non-zero entry.

It is clear that the K-SVD method uses *sparse I rank* as its optimization criterion. In this paper, we propose a new approach to address the problem in (6), using the *sparse II rank* $S_2(X)$. The problem is now formulated as

$$\left(\widehat{D}, \widehat{X}\right) = \operatorname{argmin}_{D, X} \{S_2(X) + \|Y - DX\|_F^2\} \tag{7}$$

The K-SVD method treats the columns of Y independently. By contrast, our approach considers all columns simultaneously; therefore, it is able to update multiple columns in the dictionary at each iteration. To solve the MMV problem, we propose a new MMV pursuit algorithm, called Enhanced Orthogonal Matching Pursuit or EOMP. This algorithm improves upon the the traditional OMPMMV shown in Table 1 in two ways: (i) it selects more than one atom at each iteration; (ii) it keeps a compact solution by discarding irrelevant atoms. The steps of the EOMP algorithm are summarized in Table 2 below.

Table 2. Enhanced Orthogonal Matching Pursuit Algorithm

Input
matrix $Y \in \mathbb{R}^{n \times m}$,
matrix $D \in \mathbb{R}^{n \times k}$,
maximum number of selected atoms K_{max} ,
thresholds λ and γ .
Output
matrix $X \in \mathbb{R}^{k \times m}$
Procedure
1. <i>Initialization</i>
residual error: $R_0 = Y$,
set of selected atoms: $\mathcal{S} = \emptyset$.
2. <i>Subset Selection</i>
Find atom \mathbf{d}_i in D so that $\ \mathbf{c}_i\ _q \geq \lambda \sup_{j=1, j \neq i}^k \ \mathbf{c}_j\ _q$, where $\mathbf{c}_i = (R_{t-1})^T \mathbf{d}_i$ and $q > 1$.
Add selected atoms to \mathcal{S} : $S = S \cup \mathbf{d}_i$.
3. <i>Atom Extraction</i>
3.1 Discard \mathbf{d}_i if its coherence satisfies $ \langle \mathbf{d}_i, \mathbf{d}_k \rangle \leq \gamma$, for all $\mathbf{d}_k \in \mathcal{S}$, where $i \neq k$.
3.2 If $ \mathcal{S} > K_{max}$, delete $(\mathcal{S} - K_{max})$ atoms with the lowest scores $\ \mathbf{c}_i\ _q$.
4. <i>Solution Update</i>
Find X that minimizes $\ \mathcal{S}X - Y\ _F^2$.
Update the residual error $R_t = Y - \mathcal{S}X$.
5. <i>Stopping Criterion</i>
If the number of columns in \mathcal{S} is equal to K_{max} , stop.
Otherwise, go to Step 2.

Remark 1. The thresholds λ and γ control the size of S for the orthogonal projection, and they play an important role in EOMP. Our method does not recycle the same atoms like OMPMMV does. A larger λ means more atoms will be selected at each iteration. When $\lambda = 1$, EOMP behaves like the traditional OMPMMV. Furthermore, EOMP eliminates similar atoms; the threshold γ determines how

many atoms should be discarded. In contrast, OMPMMV keeps all the atoms found, even if they are highly correlated; this tends to slow down convergence of OMPMMV. In this paper, we use $\lambda = 0.8$ and $\gamma = 0.45$.

Remark 2. Suppose that at the t -th iteration, EOMP selects m_1 atoms, but only m_2 of them are kept for the update Step, where $m_1 \geq m_2$. The complexity of EOMP is

$$O(kp_1 + m_1(m_1 - 1)/2 + m_2p_2) \quad (8)$$

where p_1 is the cost of multiplying R_{t-1} and an atom, and p_2 is the cost of *Solution Update* step. The three terms in (8) correspond to the three steps of *Subset Selection*, *Atoms Extraction*, and *Solution Update*. Subset Selection step involves k atoms in D . Atoms Extraction step needs to compare m_1 selected atoms. Solution Update step requires solving a linear equation with m_2 coefficients. By comparison, OMPMMV has, at the t -th iteration, a complexity of $O(kp_1 + m_2p_2)$. However, OMPMMV needs m_2 iterations to find m_2 atoms whereas EOMP may require only one or two iterations to locate the possible candidates.

Remark 3. For comparison purposes, we select the parameters of EOMP so that it uses the same amount of storage as the K-SVD method. The maximum number of selected atoms K_{max} can then be calculated as

$$K_{max} = \left\lfloor \frac{c \times a \times N_s + b \times a \times N_s}{c \times N_s + b} \right\rfloor \quad (9)$$

where c is number of bits required to store a coefficient, a is the maximum number of selected atoms in K-SVD, N_s is the number of columns in the target image, and b is the number of bits required to store a coefficient index.

Generally the nonzero elements in the solution obtained by EOMP are clustered in a few rows; therefore, we can update more than one atom in D at each iteration. To aid the explanation, we next define an operator called *svds*. By the singular value decomposition, a matrix $A \in \mathbb{R}^{m \times n}$ can be written as

$$A = U \Sigma V^T \quad (10)$$

where $U \in \mathbb{R}^{m \times m}$ and $V \in \mathbb{R}^{n \times n}$ are orthogonal matrices. Given a positive integer k , we define the *svds* operator as

$$svds(A, k) = \{U_k, \Sigma_k, V_k^T\} \quad (11)$$

where U_k is the first k columns in U , Σ_k is a diagonal matrix of size $k \times k$, and V_k^T is the first k columns in V^T .

The EOMP strategy for updating the dictionary D and the coefficients matrix is shown in Table 3.

Next we propose an improved K-SVD algorithm that combines the Enhanced OMP, presented in Table 2, and the Dictionary Update Strategy, presented in Table 3. This algorithm, called IK-SVD, is shown in Table 4.

Table 3. Dictionary Update Strategy

At iteration t , perform Steps 1 to 4:

1. Select N atoms from D^{t-1} that correspond to non-empty rows $\{X_s^{t-1}\}$ in X^{t-1} :

$$D_s^{t-1} = \{\mathbf{d}_i^{t-1} \in D^{t-1} \mid \text{column } \mathbf{x}_i^{t-1} \neq \mathbf{0}\}.$$
where N is a predefined integer and \mathbf{x}_i^{t-1} is the i -th row in X^{t-1} .
 2. Calculate the error due to D_s^{t-1} :

$$E_s^{t-1} = Y - \sum_{\mathbf{d}_i^{t-1} \in D_s^{t-1}} \mathbf{d}_i^{t-1} \mathbf{x}_i^{t-1}.$$
 3. Apply the *svds* operator to E_s^{t-1} to compute X_s^t and D_s^t :
 $\{U_N, \Sigma_N, V_N^T\} = \text{svds}(E_s^{t-1}, N)$, and $D_s^t = U_N$, $X_s^t = \Sigma V_N^T$.
 4. Repeat Steps 1 to 3 until all non-empty rows in X^{t-1} have been processed.
-
-

Table 4. Improved K-SVD algorithm for signal representation (IK-SVD)

Input:

- matrix Y in $\mathbb{R}^{n \times m}$,
- maximum number of selected atoms K_{max} ,
- thresholds λ and γ , and
- maximum number of iterations T_{max} .

Output:

- coefficient matrix X of size $k \times m$,
- dictionary D of size $n \times k$.

Procedure:

1. Initialize a random dictionary D^0 and coefficient matrix X_a .
 2. Repeat Steps 3 and 5 for $t = 1, 2, \dots, T_{max}$.
 3. Apply EOMP algorithm (Table 2) on $\{Y, D^{t-1}\}$ to obtain X_b .
 4. Set $X_{t-1} = X_a$ or $X_{t-1} = X_b$, to give the smallest reconstruction error:

$$\|Y - D_{t-1} X_{t-1}\|$$
 5. Apply Dictionary Update Strategy (Table 3)
on $\{Y, D^{t-1}, X^{t-1}\}$ to obtain D^t and X_a .
-
-

Remark 4. The main difference between the proposed IK-SVD algorithm and the traditional K-SVD algorithm is that a more optimal coefficient matrix is selected from the two outputs produced by the EOMP step (Step 3) and the Dictionary Update step (Step 5). This leads to better convergence for the IK-SVD.

4 Experiments and Analysis

In this section, we analyze the performance of the proposed algorithm, and compare it with the traditional K-SVD algorithm. First, we test the convergence of the proposed method using a simulated data set. Second, we evaluate the proposed algorithm and the K-SVD on an image representation task. To measure

the error between the original signal and the reconstructed signal, we use the Peak Signal-to-Noise Ratio (PSNR):

$$\text{PSNR} = 20 \log_{10}(255/\text{RMSE}), \quad (12)$$

where RMSE denotes the root-mean-square error between columns of Y and Y^* .

4.1 Convergence Analysis

A signal Y is created in the range of $[-1, 1]$. It has a dimension of 50. White Gaussian noise is added; the signal-to-noise ratio (SNR) has values of 10dB, 20dB, and 30dB. The dictionary D has initially 50 atoms, each of which is normalized to a unit l^2 -norm. The total number of training iterations is set to 30.

In this experiment, we examine the effects of three factors on the K-SVD and the proposed method: (i) the number of selected atoms K_{max} , (ii) the size of dictionary k , and (iii) the number of samples m in Y . We vary one factor while keeping the rest the same. For the K-SVD, we apply OMP algorithm to find the sparse coefficients in each column of the solution X . The number of coefficients ranges from 8 to 12. For the proposed IK-SVD, we use $\lambda = 0.8$ and $\gamma = 0.45$.

Table 5. Comparison of K-SVD and IK-SVD on simulated signals with white Gaussian noise

(K_{max}, k, m)	Execution Time (s)			PSNR (db)		
	SNR=10	SNR=20	SNR=30	SNR=10	SNR=20	SNR=30
(8, 50, 1000)						
K-SVD	46.54	46.70	47.80	66.68	69.25	70.66
IK-SVD (proposed)	10.52	10.72	10.98	68.78	71.90	72.28
(12, 50, 1000)						
K-SVD	85.96	86.48	89.09	64.11	65.47	65.74
IK-SVD (proposed)	12.35	12.24	13.11	67.16	69.60	70.36
(12, 70, 1000)						
K-SVD	92.42	92.46	92.46	64.17	65.27	65.56
IK-SVD (proposed)	15.71	16.89	17.37	64.92	65.96	66.88
(12, 70, 2000)						
K-SVD	174.52	174.64	174.76	63.69	64.91	65.25
IK-SVD (proposed)	28.03	28.43	28.79	65.18	66.36	66.91

Table 5 shows the execution time and the reconstruction error for both K-SVD and IK-SVD methods, at different noise levels. In terms of execution time, IK-SVD method runs between 3.6 and 6.9 times faster than the K-SVD. The improvement is most significant when the number of selected atoms (K_{max}) and the number of signal samples (m) are high. An explanation for this result is that, at each iteration, the proposed method can update multiple atoms in D . In addition, the IK-SVD takes less time to find an $S_2(X)$ solution compared to the traditional OMP method used in K-SVD. In terms of reconstruction error,

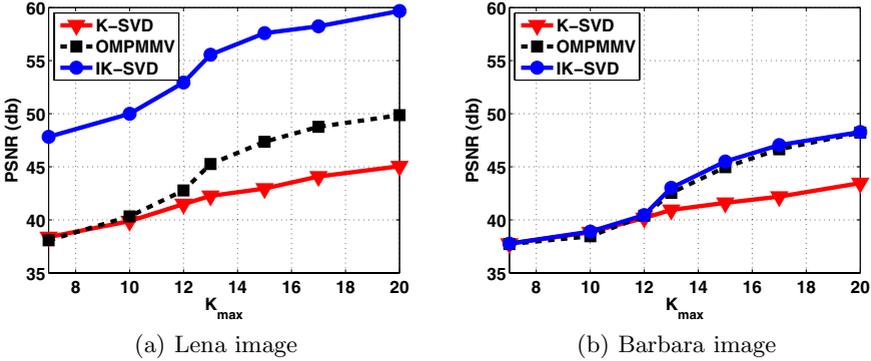


Fig. 1. Comparison of K-SVD, OMPMMV, and IK-SVD in image representation task: PSNR of reconstructed images

the IK-SVD method has higher PSNRs compared to the K-SVD in all cases. In summary, incorporation of MMV into K-SVD method improves significantly the performance of signal representation.

4.2 Application to Imaging Representation

In this experiment, we apply the IK-SVD algorithm on two images: *Barbara* and *Lena*. The size of these images is 512×512 pixels. First, each image is partitioned into non-overlapping blocks of size $M \times N$ pixels, where $M = N = 8$. A matrix Y is formed for training, by randomly selecting among these blocks; each block forms one column in Y . We apply both the K-SVD and IK-SVD methods on Y to extract two dictionaries. The dictionary size is set to $8 \times M \times N$. We also implement the OMPMMV method for comparison.

For K-SVD, we set the maximum number of atoms (K_{max}) in the range from 7 to 20. For IK-SVD and OMPMMV, the maximum number of atoms (K_{max}) is adaptively computed as in (9), so that the same amount of storage is used by EOMP and K-SVD. We run the test 20 times. The PSNR and execution time are shown in Fig. 1 and Fig. 2, respectively.

Fig. 1 and Fig. 2 show that even OMPMMV is better than that of K-SVD. The only difference between them is that OMPMMV uses $S_2(X)$, whereas K-SVD uses $S_1(X)$. This demonstrates the advantage of the MMV-based scheme. Compared to the proposed algorithm IK-SVD, OMPMMV algorithm needs more computation time to converge. This is because IK-SVD can select several atoms from D at each step. Also, the result shows that when the maximum number of atoms K_{max} increases, the accuracy of all algorithms is enhanced. However, the IK-SVD method has a lower reconstruction error than the K-SVD and OMPMMV. When K_{max} increases from 10 to 20, K-SVD execution time increases sharply, whereas IK-SVD appears more stable and robust.

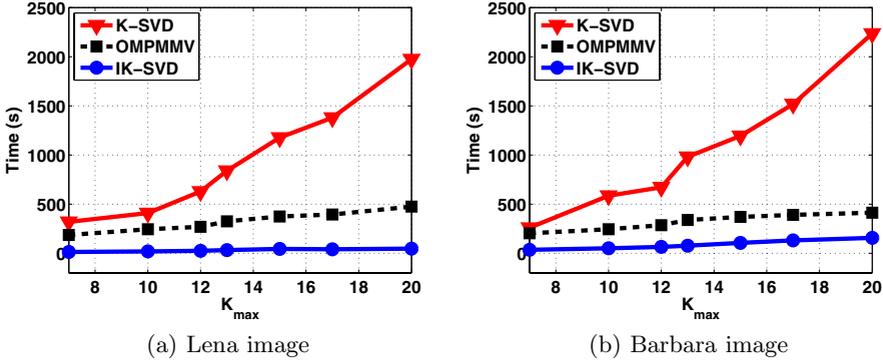


Fig. 2. Comparison of K-SVD, OMPMMV, and IK-SVD in image representation task: Execution time

5 Conclusion

This paper addresses the problem of sparse decomposition for signal representation. To date, K-SVD is the state-of-the-art method for solving this problem. We have proposed here a new method that is based on Multiple Measurement Vectors (MMV), which offer a better trade-off between computational accuracy and storage requirement. The proposed method uses an enhanced MMV pursuit algorithm (EOMP) to find a minimal- $S_2(X)$ solution and the multi-singular value decomposition to accelerate processing. Experimental results indicated that the new method runs 3.6 to 6.9 times faster, and has lower reconstruction error, compared to the existing K-SVD algorithm.

References

1. Wickerhauser, M.V.: Adapted wavelet analysis from theory to software. A. K. Peters, Ltd., Natick (1994)
2. Donoho, D.L.: On Minimum Entropy Segmentation, pp. 233–269. Academic Press, Inc., New York (1994)
3. Haupt, J., Nowak, R.: Signal Reconstruction From Noisy Random Projections. *IEEE Transactions on Information Theory* 52, 4036–4048 (2006)
4. Rauhut, H., Schnass, K., Vandergheynst, P.: Compressed Sensing and Redundant Dictionaries. *IEEE Transactions on Information Theory* 54, 2210–2219 (2008)
5. Tropp, J.A., Gilbert, A.C.: Signal Recovery From Random Measurements Via Orthogonal Matching Pursuit. *IEEE Transactions on Information Theory* 53, 4655–4666 (2007)
6. Aharon, M., Elad, M., Bruckstein, A.: K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation. *IEEE Transactions on Signal Processing* 55, 4311–4322 (2006)
7. Elad, M., Aharon, M.: Image Denoising Via Sparse and Redundant Representations Over Learned Dictionaries. *IEEE Transactions on Image Processing* 15, 3736–3745 (2006)

8. Mairal, J., Elad, M., Sapiro, G.: Sparse Representation for Color Image Restoration. *IEEE Transactions on Image Processing* 17, 53–69 (2008)
9. Cotter, S.F., Rao, B.D., Kjersti, E., Kreutz-Delgado, K.: Sparse solutions to linear inverse problems with multiple measurement vectors. *IEEE Transactions on Signal Processing* 53, 2477–2488 (2005)
10. Chen, J., Huo, X.: Theoretical Results on Sparse Representations of Multiple-Measurement Vectors. *IEEE Transactions on Signal Processing* 54, 4634–4643 (2006)
11. Rao, B.D., Engan, K., Cotter, S.F., Palmer, J., Kreutz-Delgado, K.: Subset selection in noise based on diversity measure minimization. *IEEE Transactions on Signal Processing* 51, 760–770 (2003)